DOCUMENT RESUME

ED 134 596 TM 005 982

AUTHOR Davis, Charles E.
TITLE Generalizability Theory and Bayesian Inference II.
PUB DATE Apr 76
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Bayesian Statistics; Mathematical Models
IDENTIFIERS *Bayesian Inference; *Generalizability Theory

ABSTRACT

Bayesian techniques were developed for approximating the joint and marginal distributions of universe scores in single-facit-mixed-model designs with unknown covariance structure. Empirical results indicate that when compared with Cronbach's estimates, marginal means obtained through this procedure are substantially better predictors of subsequent performance. (Author/RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED134596

Generalizability Theory and Bayesian Inference II

by

Charles E. Davis

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

[A paper presented at the AERA annual meeting in San Francisco, April 1976]

TM005 982

Davis (1975) outlined a Bayesian approach to answering some of the questions posed by generalizability theory and developed the mathematics for some rather simple models. The present discussion is intended to be a continuation of that work. As such it draws heavily upon the definitions, notation, and results developed therein.

In Davis (1975) we presumed that the covariance structures $\{C_i\}$ were known. In the present paper we intend to relax that assumption by specifying a Bayesian prior of a very specific form. In what follows we will confine our discussion to mixed-model situations and it will turn out that whenever our prior beliefs about the sampling variances $\{\sigma_t^2 , \sigma_{\tau_{*a}}^2 , \sigma_{\alpha\gamma}^2\}$ and about the hyperparameters $\{\sigma_\alpha^2 , \sigma_\mu^2\}$ may be described within a certain class, approximation techniques due to Box and Tiao (1973) can be profitably employed.

In Davis (1975) it was shown that the sampling distribution of the observations in traditional criterion-referenced testing applications may be represented within the framework of a two-stage sampling model.

$$\text{Stage I:} \quad p(\underset{\sim}{X} | \underset{\sim}{\theta}_1 , C_1) \sim N(A_1\underset{\sim}{\theta}_1 , C_1)$$

$$\text{Stage II:} \quad p(\underset{\sim}{\theta}_1 | \underset{\sim}{\theta}_2 , C_2) \sim N(A_2\underset{\sim}{\theta}_2 , C_2)$$

The parameters $\{A_i , \underset{\sim}{\theta}_i , C_i ; i = 1,2\}$ are given by the relationships in Equation (5.5) . We then added a third stage to describe the prior distribution on $\underset{\sim}{\theta}_2$ which contains the universe scores, τ_{*a} . [see equation (5.9)]

Stage III: $p(\underset{\sim}{\theta}_2 | \underset{\sim}{\theta}_3, C_3) \sim N(A_3\underset{\sim}{\theta}_3, C_3)$

It can be shown that the marginal distribution on the observations is given by:

$$p(\underset{\sim}{X} | \{A_1, \underset{\sim}{\theta}_1, C_1 ; i = 1, 2, 3\}) \sim N(A_1A_2A_3\underset{\sim}{\theta}_3, C_1 + A_1[C_2 + A_2C_3A_2^T]A_1^T)$$

with the inverse of the dispersion matrix equal to

$$\sigma_e^{-2}I - n_k^{-2}A_1[n_k\sigma_e^{-2}I - (n_k^{-2}\sigma_e^2 I + \ell_T + A_2C_3A_2^T)^{-1}]A_1^T .$$

For the special case in which ℓ_T takes the highly symmetric form

$$\ell_T = \{[n_a/(n_a - 1)\sigma_{\alpha\gamma}^2]I_{n_a} + [\sigma_{\tau_{g*}}^2 - (n_a - 1)^{-1}\sigma_{\alpha\gamma}^2]J_{n_a,n_a}\} \times I_{n_g} ,$$

this marginal distribution is given by

$$p(\underset{\sim}{X} | \theta, \sigma_e^2, \sigma_{\alpha\gamma}^2, \sigma_{\tau_{g*}}^2, \sigma_{\tau_{*a}}^2, \sigma_\mu^2) \propto \phi_e^{-\frac{\nu_e}{2}} \phi_{et}^{-\frac{\nu_t}{2}} \phi_{eta}^{-\frac{\nu_a}{2}} \phi_{eg}^{-\frac{\nu_g}{2}} \phi_{egm}^{-\frac{\nu_m}{2}}$$
$$* \exp\left\{-\frac{1}{2}\left[\frac{S_e}{\phi_e} + \frac{S_t}{\phi_{et}} + \frac{S_a}{\phi_{eta}} + \frac{S_g}{\phi_{eg}} + \frac{S_m}{\phi_{egm}}\right]\right\} \qquad (1)$$

where

$$S_m = n_g n_a n_k(\theta - X_{...})^2 \qquad \nu_m = 1$$

$$S_a = n_g n_k \Sigma (X_{.a.} - X_{...})^2 \qquad \nu_a = n_a - 1$$

$$S_g = n_a n_k \Sigma (X_{g..} - X_{...})^2 \qquad \nu_g = n_g - 1$$

$$S_t = n_k \Sigma\Sigma (X_{ga.} - X_{.a.} - X_{g..} + X_{...})^2 \qquad \nu_t = (n_a - 1)(n_g - 1)$$

$$S_e = \Sigma\Sigma\Sigma (X_{gak} - X_{ga.})^2 \qquad \nu_e = n_g n_a (n_k - 1)$$

$$\phi_{egm} = \sigma_e^2 + n_a n_k \sigma_{\tau_{g^*}}^2 + n_g n_a n_k \sigma_\mu^2$$

$$\phi_{eta} = \sigma_e^2 + n_a n_k (n_a - 1)^{-1} \sigma_{\alpha\gamma}^2 + n_g n_k n_a (n_a - 1)^{-1} \sigma_\alpha^2$$

$$\phi_{eg} = \sigma_e^2 + n_a n_k \sigma_{\tau_{g^*}}^2$$

$$\phi_{et} = \sigma_e^2 + n_a n_k (n_a - 1)^{-1} \sigma_{\alpha\gamma}^2$$

$$\phi_e = \sigma_e^2$$

Notice that if the prior were of the form

$$p(\theta, \phi_e, \phi_{et}, \phi_{eta}, \phi_{eg}, \phi_{egm}) \propto$$

$$\phi_e^{-(\eta_e + 2)/2} \phi_{et}^{-(\eta_t + 2)/2} \phi_{eta}^{-(\eta_a + 2)/2} \phi_{eg}^{-(\eta_g + 2)/2} \phi_{egm}^{-(\eta_m + 2)/2}$$

$$* \exp\left\{ -\frac{1}{2} \left[\frac{\lambda_e}{\phi_e} + \frac{\lambda_t}{\phi_{et}} + \frac{\lambda_a}{\phi_{eta}} + \frac{\lambda_g}{\phi_{eg}} + \frac{\lambda_m}{\phi_{egm}} \right] \right\} \qquad (2)$$

$$0 \le \phi_e \le \phi_{et} \le \phi_{eta}$$

$$0 \le \phi_e \le \phi_{eg} \le \phi_{egm}$$

then the posterior would be of the same form as the prior with

Prior Parameter	replaced by	Posterior Parameter
η_e		$\eta_e + \nu_e$
η_t		$\eta_t + \nu_t$
η_a		$\eta_a + \nu_a$
η_g		$\eta_g + \nu_g$
η_m		$\eta_m + \nu_m$
λ_e		$\lambda_e + S_e$
λ_t		$\lambda_t + S_t$
λ_a		$\lambda_a + S_a$
λ_g		$\lambda_g + S_g$
λ_m		$\lambda_m + n_g n_a n_k (\theta - x_{...})^2$

Thus as long as a member of the family of priors described by Equation (2) can be found which provides an adequate description of our beliefs in any given situation, we will take the position that convenience dictates its use. But specifying a prior of this form is no small task for it implies that knowledge about each of the parameters ϕ_e ,ϕ_{et} ,ϕ_{eta} ,ϕ_{eg} , and ϕ_{egm} is correlated with knowledge about each of the others. The task would be somewhat simpler if ϕ_e were known apriori. In that case our researcher could specify joint priors on each of the pairs $\{\phi_{et}$,$\phi_{eta}\}$ and $\{\phi_{eg}$,$\phi_{egm}\}$ independently. But in the usual case ϕ_e is not known, and therefore, knowledge of each of these parameters is correlated. However, if

considerable information is available on ϕ_e, one might obtain useful approximations to the priors on (ϕ_{et}, ϕ_{eta}) and (ϕ_{eg}, ϕ_{egm}) by assessing priors assuming ϕ_e is equal to its prior marginal mode. Then provided the general shapes of these distributions remain constant throughout the range of ϕ_e, the estimates of η_t and λ_t, say obtained using the marginal mode of ϕ_e would be the same as the η_t and λ_t in the conditional distributions for general ϕ_e. If we are to use this procedure it must be true that the parameters describing the conditional priors on (ϕ_{et}, ϕ_{eta}) and (ϕ_{eg}, ϕ_{egm}) do not change radically with changes in ϕ_e throughout its probable range; only the point of truncation of these members changes. This formulation is likely to hold only in those cases in which there is considerable prior information about ϕ_e so that its probable range is severely restricted.

In much of the remainder of this paper we shall be dealing with random variables Y which will be distributed such that $Z = (\Delta + \kappa Y)/\lambda$ has an inverse chi-square distribution on η degrees of freedom, where Δ, κ, and λ are constants. For convenience we provide the following reduced notation.

Definition:

A variable Y will be said to have an inverse chi-square distribution with parameters η, λ, κ, and Δ and will be denoted $\chi^{-2}(\eta, \lambda, \kappa, \Delta)$, if $Z = (\Delta + \kappa Y)/\lambda$ has a standard inverse chi-square distribution on η degrees of freedom.

Before proceeding further, it will be necessary to state an important result due to Box and Tiao (1973). If the probability density of two variables σ_1^2 and σ_2^2 is such that

$$p(\sigma_1^2,\sigma_2^2) \propto (\sigma_1^2)^{-(\frac{1}{2}n_1+1)} (\sigma_1^2+\kappa\sigma_2^2)^{-(\frac{1}{2}n_2+1)} \exp\left\{ -\tfrac{1}{2}\left(\frac{\lambda_1}{\sigma_1^2} + \frac{\lambda_2}{\sigma_1^2+\kappa\sigma_2^2} \right) \right\}$$

for n_1, n_2, λ_1, and λ_2 positive constants and for $\sigma_1^2 > 0$ and $\sigma_2^2 > 0$, then the moment generating function for $W = \lambda_1/\sigma_1^2$ is given by

$$M_W(t) = (1-2t)^{-\frac{n_1}{2}} \frac{\Pr\left\{ F_{n_2,n_1} < \frac{\lambda_2 n_1/(n_2\lambda_1)}{1-2t} \right\}}{\Pr\left\{ F_{n_2,n_1} < \lambda_2 n_1/(n_2\lambda_1) \right\}}$$

for $|t| < \frac{1}{2}$.

By equating the first two moments of W to that of a chi-square variable, with degrees of freedom b and scale parameter a, $\chi^2(b, a)$, we find

$$a = \left(\frac{n_1}{2} + 1 \right) \frac{I_x(\frac{1}{2}n_2, \frac{1}{2}n_1+2)}{I_x(\frac{1}{2}n_2, \frac{1}{2}n_1+1)} - \frac{n_1}{2} \frac{I_x(\frac{1}{2}n_2, \frac{1}{2}n_1+1)}{I_x(\frac{1}{2}n_2, \frac{1}{2}n_1)}$$

$$b = \frac{n_1}{a} \frac{I_x(\frac{1}{2}n_2, \frac{1}{2}n_1+1)}{I_x(\frac{1}{2}n_2, \frac{1}{2}n_1)}$$

where

$$x = \frac{\lambda_2}{\lambda_1 + \lambda_2} .$$

Therefore, to this degree of approximation,

$$p(\sigma_1^2) \doteq \chi^{-2}(\eta_1^* = b,\ \lambda_1^* = (\lambda_1/a),\ 1,\ 0) \ .$$

Next, we define a new variable $\sigma_{12}^2 = \sigma_1^2 + \kappa\sigma_2^2$. A change of variable in Equation (3.1.11) yields

$$p(\sigma_1^2,\ \sigma_{12}^2) \propto (\sigma_1^2)^{-(\frac{1}{2}\eta_1+1)} (\sigma_{12}^2)^{-(\frac{1}{2}\eta_2+1)} \exp\left\{ -\tfrac{1}{2}\left[\frac{\lambda_1}{\sigma_1^2} + \frac{\lambda_2}{\sigma_{12}^2}\right]\right\}$$

$$\sigma_1^2 > 0 \text{ and } \sigma_{12}^2 > \sigma_1^2 \ .$$

The rth moment for $W_1 = \lambda_2/\sigma_{12}^2$ is given by

$$\mathcal{E}[W_1^r] = \left(\frac{2}{\lambda_2}\right)^r \frac{\Gamma\left(\frac{\eta_2}{2}+r\right)}{\Gamma\left(\frac{\eta_2}{2}\right)} \quad \frac{I_x\left(\frac{\eta_2}{2}+r,\ \frac{\eta_1}{2}\right)}{I_x\left(\frac{\eta_2}{2},\ \frac{\eta_1}{2}\right)}$$

where

$$x = \frac{\lambda_2}{\lambda_1 + \lambda_2} \ .$$

Equating the first two moments of W_1 to those of a chi-square variable with scale parameter c and degrees of freedom d, $\chi^2(d,\ c)$, we find

$$c = \left(\frac{\eta_2}{2} + 1\right) \frac{I_x(\frac{1}{2}\eta_2+2, \frac{1}{2}\eta_1)}{I_x(\frac{1}{2}\eta_2+1, \frac{1}{2}\eta_1)} - \frac{\eta_2}{2} \frac{I_x(\frac{1}{2}\eta_2+1, \frac{1}{2}\eta_1)}{I_x(\frac{1}{2}\eta_2, \frac{1}{2}\eta_1)}$$

and $$d = \frac{\eta_2}{2} \frac{I_x(\frac{1}{2}\eta_2+1, \frac{1}{2}\eta_1)}{I_x(\frac{1}{2}\eta_2, \frac{1}{2}\eta_1)} .$$

Therefore, to this degree of approximation,

$$p(\sigma_{12}^2) \doteq \chi^{-2}(\eta_2^* = d,\ \lambda_2^* = \lambda_2/c,\ 1,\ 0) .$$

What we are proposing and what we hope to report on shortly is the following.

1. Specify a marginal on σ_e^2 by selecting a member of the general inverse chi-square class $\chi^{-2}(\eta_e^*, \lambda_e^*, 1, 0)$ by applying Schlaifer's (1971) fractile assessment technique.
2. Assuming σ_e^2 equals its prior mode $\hat{\sigma}_e^2$, use a modified form of Schlaifer's technique to specify inverse chi-square priors, $\chi^{-2}(\eta_t^*, \lambda_t^*, n_a n_k (n_a - 1)^{-1}, \hat{\sigma}_e^2)$ and $\chi^{-2}(\eta_g^*, \lambda_g^*, n_a n_k, \hat{\sigma}_e^2)$, on σ_t^2 and $\sigma_{\tau_{g^*}}^2$, respectively.
3. Assuming $(\sigma_e^2, \sigma_t^2, \sigma_{\tau_{g^*}}^2)$ equals the joint mode $(\hat{\sigma}_e^2, \hat{\sigma}_t^2, \hat{\sigma}_{\tau_{g^*}}^2)$, use Schlaifer's technique to specify priors, $\chi^{-2}(\eta_a, \lambda_a, n_g n_a n_k (n_a - 1)^{-1}, \hat{\sigma}_e^2 + n_k n_a (n_a - 1)^{-1} \hat{\sigma}_t^2)$ and $\chi^{-2}(\eta_m, \lambda_m, n_g n_a n_k, \hat{\sigma}_e^2 + n_a n_k \hat{\sigma}_{\tau_{g^*}}^2)$, on $\sigma_{\tau_{a^*}}^2$ and σ_m^2, respectively.
4. Iteratively determine λ_t and η_t so that the conditional distributions of $(\phi_{et} | \phi_e = \hat{\phi}_e)$ has parameters λ_t^* and η_t^*.
5. Iteratively determine η_g and λ_g so that the conditional distribution of $(\phi_{eg} | \phi_e = \hat{\phi}_e)$ has parameters η_g^* and λ_g^*.

6. Iteratively determine η_e and λ_e so that the marginal distributions of σ_e^2 has parameters η_e^* and λ_e^*.

7. Using Box and Tiao approximations to the marginal distributions of each parameter, verify that these priors adequately reflect your prior beliefs.

REFERENCES

Box, G. E. P., and Tiao, G. C. Bayesian Inference in Statistical Analysis. Reading Mass.: Addison-Wesley Publishing Company, 1973.

Davis, C. E. Generalizability theory and Bayesian Inference. A paper presented at the annual meeting of the American Educational Research Association in Washington, D. C. April 1975.

Schlaifer, R. Computer Programs for Elementary Decision Analysis. Division of Research, Graduate School of Business Administration. Boston: Harvard University, 1971.